AD/A-001 813

SOFTWARE DEVELOPMENT FOR DECISION ANALYSIS

Allen C. Miller, III

Stanford Research Institute

Prepared for:

Defense Supply Service
Defense Contract Administration Services Region
Defense Advanced Research Projects Agency

November 1974

DISTRIBUTED BY:

NTIS
National Technical Information Service
U. S. DEPARTMENT OF COMMERCE

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER AD/A001813
4. TITLE (and Subtitle) Software Development for Decision Analysis		5. TYPE OF REPORT & PERIOD COVERED Semi-Annual Technical Report 1 March 1974-1 November 1974
		6. PERFORMING ORG. REPORT NUMBER SRI-3309
7. AUTHOR(s) Allen C. Miller, III		8. CONTRACT OR GRANT NUMBER(s) MDA903-74-C-0240
9. PERFORMING ORGANIZATION NAME AND ADDRESS Stanford Research Institute 333 Ravenswood Avenue Menlo Park, California 94025		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
11. CONTROLLING OFFICE NAME AND ADDRESS Defense Supply Service The Pentagon Washington, D. C.		12. REPORT DATE November 1974
		13. NUMBER OF PAGES 13
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office) Defense Contract Administration Services Region, San Francisco 866 Malcolm Road Burlingame, California 94010		15. SECURITY CLASS. (of this report) Unclassified
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Unlimited Distribution

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

Unlimited Distribution

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

i

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

Unclassified

SEMI-ANNUAL TECHNICAL REPORT

SOFTWARE DEVELOPMENT FOR DECISION ANALYSIS

Contract Number: MDA903-74-C-0240

DARPA Order Number: 2742

Program Code Number: PD420

Contractor: Stanford Research Institute
333 Ravenswood Avenue
Menlo Park, California 94025

Effective Date of Contract: 1 June 1974

Contract Expiration Date: 30 September 1975

Amount of Contract: $155,000

Principal Investigator: James E. Matheson
(415) 326-6200, ext. 4450

Project Scientist: Allen C. Miller, III
(415) 326-6200, ext. 2865

SRI Project Number: 3309

SRI Proposal Number: MSU-73-133

Sponsored by

Defense Advanced Research Projects Agency

DARPA Order Number 2742

The views and conclusions contained in this document are those of the authors and should not be interpreted as necessarily representing the official policies, either expressed or implied, of the Defense Advanced Research Projects Agency or the U. S. Government

SEMI-ANNUAL TECHNICAL REPORT

SOFTWARE DEVELOPMENT FOR DECISION ANALYSIS

I. Summary

This report describes the research carried out by the SRI Decision Analysis Group for the Defense Advanced Research Projects Agency (DARPA) under contract number MDA903-74-C-0240 during the period 1 March 1974 to 1 November 1974. This research effort has two objectives:

Task A - Develop a morphology for characterizing and analyzing decision problems to serve as a basis for the design of a system of integrated computer aids for decision making.

Task B - Transfer the existing SRI CTREE program to two computers, one available for classified work in the Washington, D. C. area and one accessible through the DARPA computer network.

The work in Task A has resulted in a definition of the operational components of a decision morphology. Some of these components have been studied in preliminary research efforts designed to determine which areas should be examined in more detail. Preliminary research has been conducted on eight of the fourteen decision-morphology components currently defined. The results of these research projects have been summarized in five internal technical memoranda and a tabular summary of past decision analysis projects. In addition, a conference paper has been presented on one of the morphology components, and another paper is being prepared for publication. Additional technical memoranda are currently in preparation.

As part of Task B, the SRI CTREE program has been installed on the UCLA-CCBS PDP-10 computer. This computer can be accessed through the DARPA computer network. The CTREE programs have been run and verified by SRI personnel. A user's manual for the software has been written and is currently undergoing final editing. DARPA has not yet selected the second computer site in the Washington, D. C. area, and has not established a firm date for a training session for new users of the software.

II. Task A: Development of a Decision Morphology

Components of a Decision Morphology

Following the research plan outlined in SRI Proposal No. MSU-73-133, we have broken the morphology development effort down into the set of operational components listed in Table 1. This list is tentative and will probably be modified as the research develops. The components in Table 1 are not necessarily of equal importance to a decision analysis problem; in fact, some of the areas may not be needed for the analysis of a simple problem. However it is necessary to deal with all of them to develop an integrated set of decision aids, since different morphology components are needed for different types of decision problems.

Since analysts use different approaches to structuring decision problems, there has been considerable discussion within the SRI Decision Analysis Group about which components of the morphology should be studied in most detail because of their practical importance. We initiated a series of preliminary research projects designed to answer this question. It is our intention to address all of the components so there will be no gaps in our understanding of the decision analysis process. However, we plan to concentrate our efforts on those components that will find application in the greatest number of decision problems.

Research Projects

Research is currently underway on languages for deterministic models, descriptions of probabilistic dependence, coalescence, and the value of decision-dependent information (items 7, 9, 10, and 11 in Table 1). Some preliminary work has also been done on an inventory of decision problems, management of model growth, languages for probabilistic processing in decision trees, and special forms of probabilistic processing (items 1, 8, 13, and 14 in Table 1). In each of these efforts we are attempting to define the relationships between the various components of the decision morphology and the manner in which the components will be combined to form an integrated set of decision aids.

A brief description of each of the research projects follows.

TABLE 1

Components of a Decision Morphology

I. Types of decision problems and associated decision aids

   1. Inventory of decision problems and classification by the types of decision aids and models appropriate to each; demonstration of each type of decision problem with a simple textbook problem

II. Elicitation of subjective information

   2. Probability encoding

      A. Individual

      B. Group

   3. Value encoding and synthesis of multi-attribute values

   4. Encoding of risk attitude

   5. Generation of alternatives

   6. Model elicitation and initial problem structuring

III. Problem structuring and modeling

   7. Algebraic and graphical languages for detemrministic models

      A. One-shot decisions vs. sequential decisions

      B. Continuous vs. discrete value functions and decision variables

   8. Management of model growth

IV. Theoretical problems associated with probabilistic models

   9. Probabilistic dependence

   10. Coalescence

   11. Value of decision-dependent information

   12. Value of sequential information

V. Analyzing probabilistic models

   13. Decision trees

      A. Description and generation of large trees

      B. Probability trees and Bayes' Rule processing

      C. Visualizing and manipulating portions of a decision tree

      D. Efficient solution techniques

   14. Other forms of probabilistic processing

      A. Monte Carlo methods

      B. Approximate methods

A. Inventory of Decision Problems: Past experience with the analysis of decision problems has shown that certain phases of the analysis are more difficult and time-consuming than others. A review of over fifty decision analysis projects carried out at SRI shows that, in general, the analysts consider such areas as problem structuring and the evaluation of complex models to be the most difficult and the most likely to benefit from an automated decision aid. However, more research is needed to determine the characteristics of decision problems that would make automated decision aids especially useful for certain portions of the analysis. A table ranking the potential usefulness of various types of decision aids for the SRI decision analysis projects has been prepared. It will form the basis of a more detailed examination of the types of problems that could benefit most from automated decision aids.

B. Management of Model Growth: Our initial efforts in this area have shown that models of a complicated decision problem do not evolve in a straightforward manner. Instead, the complexity of a model increases for a while as the analyst attempts to deal with all of the variables that might have an important impact on the decision, and then the model shrinks as preliminary calculations and sensitivity analyses are used to eliminate relatively unimportant parameters. This process of expansion and contraction is often repeated several times before the model is solved in detail; different types of decision aids are needed for the expansion and contraction phases. The interaction between the expansion and contraction phases of the analysis are discussed in an internal technical memorandum by T. Rice.(1) One of the topics that we will consider in further research in this area is the linkage that will be needed to allow the user to switch back and forth between the decision aids appropriate to each phase.

C. Languages for Deterministic Models: The most difficult and time-consuming part of a decision analysis is often the development of a deterministic model that captures the important elements of the problem. The resulting model can be simple in form, but is the result of considerable effort to determine which variables have the greatest impact on the decision and how the variables are related. These variables are selected by testing

the sensitivity of the decision to the variables as the model is developed through a process of expansion and contraction. (See the preceding discussion.) An internal technical memorandum has been prepared showing that a generalized branching structure can be used both to describe the relationships among the variables and to decompose aggregate parameters into more basic variables.[2] More research is needed to understand how the branching structures could be generated and manipulated by an automated decision aid, but it appears that such an aid would be especially useful in the expansion phase of model development. Of special concern is the question of how to change a static model to a dynamic one while continuing to make use of the previously-defined structural relationships.

D. Coalescence: Coalescence is a process of eliminating redundant portions of decision trees based on the dependencies among the uncertain variables represented in the tree and the different information states that occur at different nodes in the tree. Coalescence is the key to solving large, complex decision problems that would require a prohibitive amount of probability encoding and processing if redundancies were not eliminated. For example, a Markov process must be represented by an infinitely large decision tree if coalescence is not used to simplify the analysis. The solution of such a tree is quite difficult, but Markov processes are relatively easy to analyze when the redundancies are removed from the tree. Graphical techniques have been developed for exploiting coalescence, and these have been documented in an internal technical memorandum by A. Miller.[3]

E. Descriptions of Probabilistic Dependence: A thorough understanding of the dependencies and interdepencies inherent in assessed probabilities is necessary for both the accurate elicitation of probabilities and the efficient solution of large decision trees. Although there are many ways to elicit dependent probabilities, some are easier for the subject to think about than others, especially when the uncertainties relate to several events that could occur in any order. The use of coalescence (see the preceding discussion) to simplify large decision trees depends on a careful examination of the dependencies among random variables. These dependencies can be very complex. Two uncertain variables can be either dependent or independent. However, our research of the last few months has shown that

three variables can be statistically related in 18 different ways, ranging from mutual independence to complete independence. When there are more than three uncertain quantities, the number of possible dependencies becomes quite large. An internal technical memorandum describing these forms of dependency is currently in preparation.

F. Value of Decision-Dependent Information: When the probability associated with an uncertain quantity depends on which alternative the decision maker has chosen, the value of information about the uncertain quantity also depends on the alternative chosen, even though the information is supplied before the alternative is selected. This fact has caused us to reassess our understanding of the value of information and the process that is used to determine such a value. We now believe that a single uncertain quantity must be represented by more than one random variable when the uncertainty depends on a decision. A preliminary paper on this subject was presented at the ORSA-TIMS Conference in Puerto Rico on October 18, 1974 by J. Matheson. A more detailed paper discussing the problem and the procedures for dealing with it is currently in preparation.

G. Languages for Probabilistic Processing in Decision Trees: The current computer programs for analyzing decision trees force the user to numerically or algebraically calculate the probabilities imbedded in the tree[4,5] Some of this probabilistic processing remains the subject of research. (See the discussions of probabilistic dependence, coalescence, and the value of decision-dependent information.) However, the probabilities for most simple decision problems can be determined with Bayes' Rule. A simple language has been defined to specify uniquely the Bayes' Rule processing that should be carried out in the next generation of advanced decision tree programs. This language allows the user to input probabilities in a form that is easy to assess, and then specifies how the probabilities should be processed before they are used in the decision tree. However, the language requires a high level of competence on the part of the user. We have not yet reached the goal of a simple, direct specification of the

necessary probabilistic processing that can be used by a non-technical decision maker. The research to date on this topic has been documented in an internal technical memorandum by J. Pezier.(6)

H. Special Forms of Probabilistic Processing: For certain types of decision models, it is possible to use approximate techniques to determine the optimal decision and the value of information.(7,8) While these techniques may introduce a slight error into the solution, they make it possible to solve a very large, complex model with a minimum of effort. Furthermore, they can be used to explore efficiently the implications of the model while it is evolving, thus helping the analyst develop a model that more accurately represents the important elements of the problem. Some of these methods may prove more useful than sensitivity analysis in determining the impact of changes in uncertain quantities on the decision. While considerable theoretical work on this topic has been documented, the development of a systematic methodology for using approximate techniques to guide the analyst has only begun. The procedures for implementing some of the approximate techniques have been documented in an internal technical memorandum by T. Rice.(9)

III. Task B - Software Transfer

Background

In April of this year, preliminary work was started toward the identification of an initial computer facility upon which to make the first software installation. The UCLA-CCBS site was selected based on the following reasons:

- Only a limited number of DARPA sites supported the FORTRAN language; the language of the CTREE system.
- Of the sites that supported FORTRAN there were only two choices of hardware, IBM and PDP equipment. Because of the difficulties of converting to IBM equipment, PDP was chosen.
- Of the PDP computers available, the UCLA-CCBS system was chosen since it was scheduled to be the base system for a large portion of the decision oriented packages to be used by the ARPA community.

Installation of CTREE at UCLA

The conversion was begun in May and the CTREE system was thoroughly checked out by SRI personnel by the end of July. In the process of converting to the UCLA-CCBS system several new features and system enhancements were made to the CTREE package. These changes were made to aid the user in the generation and evaluation of large trees--areas that had been troublesome in the past.

Work on this task was suspended pending the selection of the second computer site for the software package. An attempt is being made by DARPA to locate a secure facility for the second conversion that would still give adequate accessibility to a wide range of users in the defense community. A suitable facility has been hard to find due to the problems inherent in accessing computers with classified data.

Tasks to be Done

The printing of the user's manual for the CTREE system is scheduled to be completed in December 1974. The draft material is currently in the final editing and preparation stages.

A second site conversion is yet to be done. Because of the difficulties involved in transfering the software to dissimilar computers, it would be best to choose another PDP machine for the second installation of the programs. This would minimize the time and effort required to complete the second software transfer. By working with a machine that is similar to the one at UCLA-CCBS, much of the conversion will be restricted to adjusting the present package for minor system differences, principally in the input/output areas. However, if the requirement for a secure facility is relaxed, a second system may not be needed since the UCLA-CCBS system is available to all DARPA users. If this is the case, additional training sessions could be substituted for the second software transfer.

The final task to be performed is a training session for potential users at a site selected by DARPA. At present, a letter has been mailed to Dr. Jerry Shure of the UCLA-CCBS facility asking him to forward to DARPA a list of people he would like to have attend a training session. Upon receipt of this list, DARPA is expected to designate the date, location, and attendees for the training session.

When the above items have been completed or resolved, the objectives of Task B will be accomplished and a final report on this subtask will be issued.

REFERENCES

(1)
Rice, T. R., "Managing the Growth of Models", Internal Technical Memorandum, July 1974, Stanford Research Institute.

(2)
Rice, T. R., "General Branching Structures", Internal Technical Memorandum, August 1974, Stanford Research Institute.

(3)
Miller, A. C., "Coalescence and Influence Diagrams", Internal Technical Memorandum, October 1974, Stanford Research Institute.

(4)
Rousseau, W., and North, D. W., "A User's Manual for TREE, A Program for the Analysis of Decision and Probability Trees", SRI Project 8106-6, 1971, Stanford Research Institute, Menlo Park, California.

(5)
Rousseau, W., and Zamora, R., "A User's Manual for the Generation and Analysis of Probability and Decision Trees", SRI Project 464531-307, 1971, Stanford Research Institute, Menlo Park, California.

(6)
Pezier, J., "A Language for Bayes' Rule Processing", Internal Technical Memorandum, October 1974, Stanford Research Institute.

(7)
Howard, Ronald A., "Proximal Decision Analysis", Management Science, Vol. 17, No. 9, May 1971.

(8)
Rice, Thomas R., "The Economics of Decision Making", Research Report No. EES-DA-74-1, August 31, 1974, Dept. of Engineering-Economic Systems, Stanford University, Stanford, California 94305.

(9)
Rice, T. R. , "Computerized Model Processing", Internal Technical Memorandum, September 1974, Stanford Research Institute.